Chapter 2

The Framework, Design Issues and Applications of Digital Watermarking

This chapter gives an overview of the watermarking systems. The general watermarking system framework and its basic components are described in Section 2.1. Section 2.2 describes the major design issues concerning watermarking techniques. Finally, the watermarking techniques designed for different application purposes are outlined in Section 2.3.

2.1 A General Watermarking System Framework

A watermarking system model can be viewed from a communication problem perspective because the two systems are similar in many ways. However, there are still differences. Figure 2.1 gives a general watermarking system framework, where the details of the two basic components, the watermark embedding scheme and watermark detection scheme, are shown in Figs. 2.2 and 2.3, respectively. The functional diagram and the basic elements of a digital communication system are also given in Fig. 2.4 [4]. We will describe the functions of the blocks in these figures in the next several paragraphs.

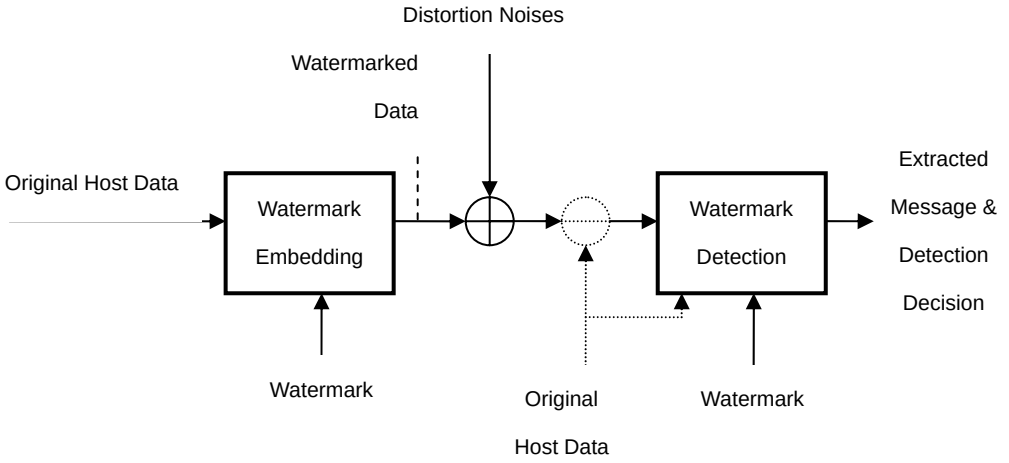


Fig. 2.1. A general watermarking system framework.

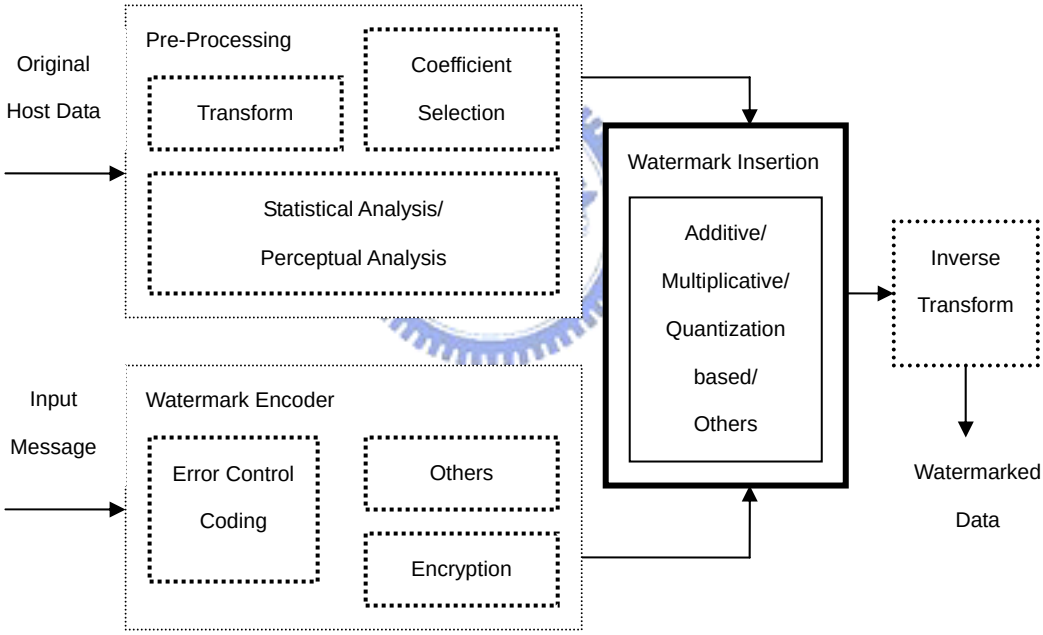


Fig. 2.2. A general watermark embedding scheme.

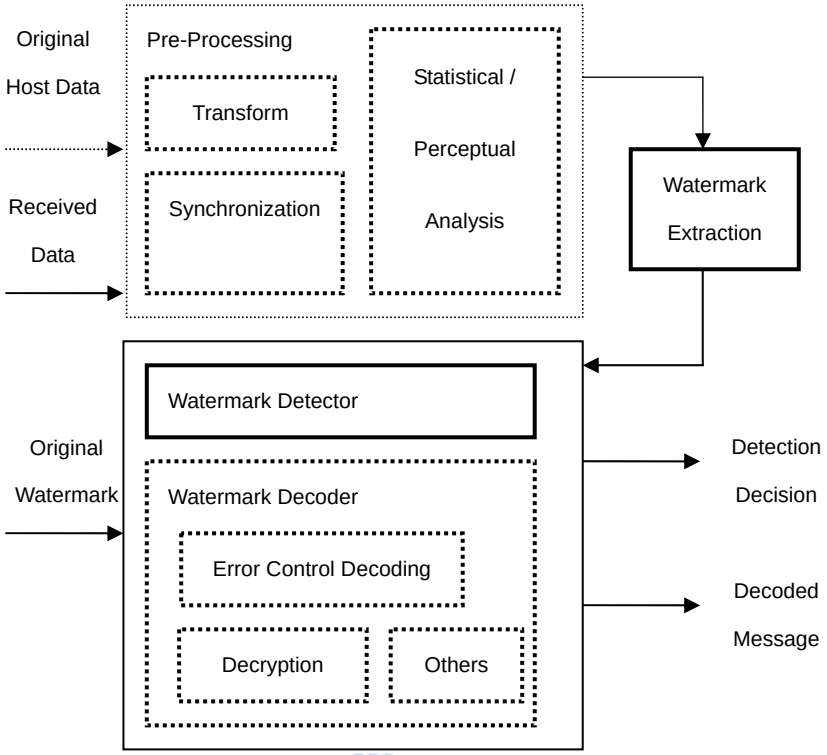


Fig. 2.3. A general watermark detection scheme.

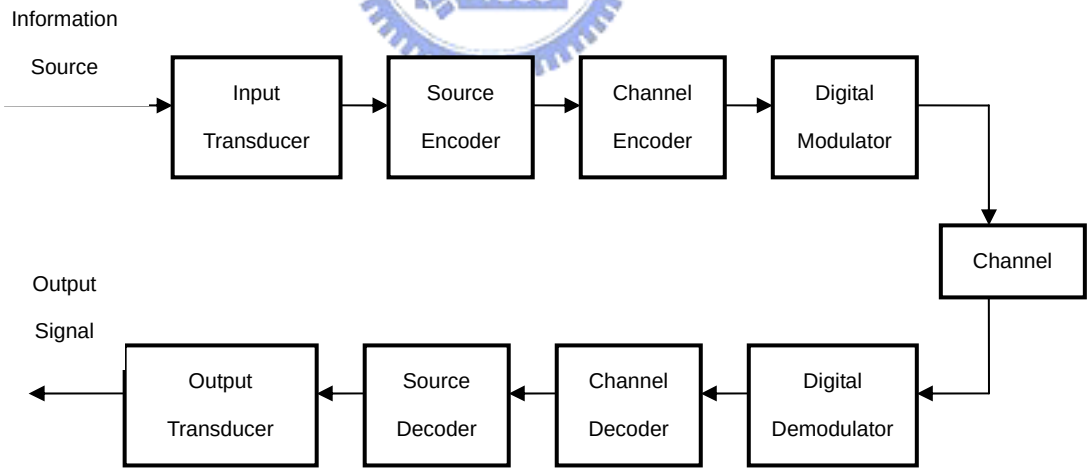


Fig. 2.4. Basic elements of a digital communication system.

The watermark is inserted in the either spatial domain or transform domains. Often, the *transform-domain watermarking* is preferred for several reasons. For example, some

transform coefficients have the good property of geometric distortion resistance. It is also easier to impose the invisibility constraints in the transform domains (e.g., Fourier transform domain and discrete cosine transform domain) since the human perceptual models (e.g. HVS, HAS) are well developed in these domains. Nowadays, popular compression algorithms are also built based on the energy-preserving transforms. Therefore, the quantization distortions imposed on the embedded watermark are easily pre-estimated if the watermark is embedded in the same coefficient representation. Finally, the previous works showed that high watermark capacity is achievable in the transform domains rather than in the spatial domain.

In a communication system, the transmitted information sequences consist of voice (speech), data (audio, video, images, text), and control messages. The source coding and channel coding techniques are applied to reduce the data redundancy and improve the data robustness. In a watermarking system, the watermark message may be meaningful information or recognizable patterns. Before watermark insertion, this message can be optionally passed through a *watermark encoder* to improve the watermark security and robustness. The security can be promoted via data encryption, where the secret key is known to both the watermark embedder and watermark detector. The error detection and error correction capabilities can be improved by employing the error control coding techniques (e.g., RS code, BCH code). The encoding technique can improve the watermark detection performance if it is well designed to match the channel characteristics (possible attacks). Note that the induced data redundancy benefits the watermark security/robustness while there is also the tradeoff between watermark capacity and watermark security/robustness.

A better watermarking scheme can be achieved by *selecting appropriate parameters* (e.g., coefficient locations, watermark strength) on the embedder side. If the watermark detector and the attack types are known in advance, the coefficients leading to poor detection performance can be pre-estimated and dropped. Note that there is tradeoff between different criterions. For example, the watermark robustness can be improved by adjusting the

watermark strength with the aid of the human perceptual models such that the watermark transparency is also retained.

The *watermark insertion* process plays the role analogous to the digital modulation process on the transmitter side in the digital communication system. A digital modulator (e.g., DPSK, *M*-ary PSK, QAM) maps a binary information sequence into an electrical waveform. In selecting a proper modulator, the probability of error is the main concern. The watermark insertion process maps the binary information sequence into the carrier (digital host data) using either additive, multiplicative, quantization based or others techniques.

The embedded watermark may be distorted before it being detected. In a communication system, the channel model is then modeled as the additive noise, linear filter channel with limited bandwidth, or the time-variant filter channel (fading multipath). Nevertheless, for the watermarking system, the channel noise is most often modeled additively for the theoretical analysis although it is not always proper. For different types of attacks, different distortion models are needed for estimating watermark capacity, error detection probability and robustness analysis. The watermark detection performance can be improved if the attack distortion information is available to the watermark embedder. Note that for blind detection, the original host data is another source of the channel noises.

Before watermark extraction, a watermark detector has to synchronize the watermark locations with those at embedding. The original host could be available (oblivious detection) or not (blind detection) on the watermark detector side. The synchronization problem is easily solved if non-blind detection is used. The optimal detection strategy is related to how the watermark is embedded. The interference problem due to host data can be also eliminated via non-blind detection. The use of blind detection or non-blind detection is application dependent.

In a communication system, an optimal receiver is designed for various modulation techniques and channel models. A *digital demodulator* converts the channel corrupted

waveform into the estimated information sequence. The *channel decoder* detects and corrects data errors while the *source decoder* decompresses and reconstructs the bitstreams. In a watermarking system, the *watermark extraction* applies the inverse operations to the received data to retrieve the embedded message. If the watermark encoder is used on the embedder side, a *watermark decoder* then decodes the extracted sequence. For some applications, a watermark detector may decode the watermark message if its existence is asserted, but others only determine if the watermark exists or not.

The similarity of the basic ideas between the communication systems and the watermark systems lead to the derivation of many watermarking theoretical works that analyze the system performance in a way similar to that in the communication systems.

2.2 Design Issues

Watermarking scheme design are application dependent. The relative importance of the requirements varies with applications. These requirements are strongly related, and thus, there are tradeoffs among them. For instance, a high watermark strength or data capacity may improve watermark robustness while it is also possible to degrade data fidelity. In the section, several general issues are addressed.

2.2.1 Robustness

The watermark robustness is measured by whether the embedded bits can be correctly decoded even when distortions occur. There is no known no watermarking scheme that can survive all types of attacks. A reasonable compromise is that the watermark robustness level is designed for certain applications and data types. For example, for media content integrity control applications, the watermark only has to fight mild distortions. The tamper indication is achieved by designing the watermark to be sensitive to sever attacks.

Image and video watermarks do not have to resist sever geometric distortions and signal processings since these attacks often lower the picture quality and thus reduce commercial value. However, these watermarks have to be robust to the attacks which remove the invisible watermarks. One example of the mild distortion is the geometric distortions with small angle rotation and small portion of cropping, which will not cause the perceptual difference between the embedded host data and the attacked host data. Another example is the data compression techniques since host data are usually compressed to reduce the storage space and transmission bit rate. For some common signal processing methods such as noise removal of the images, pitch shifting and time fluctuation of the audio signals, watermarks are expected to be able to survive through these process.

For the applications such as copyright protection, it is desirable that the watermark is detectable in a short but semantically meaningful segment of audios and videos. If the possible attacks are pre-known, the watermark can be designed more effectively and robustly.

Note that unlike the case of data encryption, attackers may break watermarking schemes without knowing the secret key for watermark embedding. For copyright protection applications, attackers are not aiming at extracting the watermark, but they just like to render the watermark undetectable while the image fidelity constraint is satisfied. This could be done by decreasing the watermark strength. An example is the collusion attack. There are two types of collusion attack. One is that several different host contents are embedded with the same watermark. The other is that several different watermarked versions are created for the same host content. For both types of attacks, attackers can estimate the watermark using a few several statistical operations.

2.2.2 Imperceptibility

Imperceptibility refers that the watermark is embedded without changing the perceptual quality of the original media. That is, the watermarked data perceptually resembles the unwatermarked one. There is a tradeoff between the watermark imperceptibility and watermark robustness. To improve the watermark robustness, the watermark strength should be maximized, which generally lower perceptual quality.

The most popular technique used to solve the above problem is employing the human perceptual models. The human perceptual sensitivity varies with individuals, and there have been several human perceptual models developed by adjusting the curve to fit to the empirical data from subjective experiments. These models are usually developed in the DCT domain, DFT domain and DWT domain. Despite these models built in different domains, in general, people are more sensitive to low frequency components than high frequency components. However, since the power of low frequency components is usually larger than the high frequency ones, it turns out that the low frequency components allow larger modification without being sensed.

2.2.3 Capacity

Watermark capacity refers to how many watermark bits can be embedded into the given host data. It is host data dependent. Similar to the capacity analysis for a communication system, the watermark capacity varies with channel. For those applications where a watermark detector only judges whether the watermark exists, the effective watermark capacity is one bit.

2.2.4 Detection Reliability

Improving the detection reliability implies decreasing the error detection probability. The

error detection probability includes both the false positive probability (false alarm) and false negative (missing) probability. The false positive probability refers to that an unwatermarked data is wrongly declared watermarked by the detector. Conversely, the false negative probability means the watermark is undetected in the watermarked data. The average error detection probability is the average of the false positive probability and the false negative probability if we assume the data is equally likely marked or unmarked. As to how small the error detection probability should be, it is application dependent. Note that more watermarked coefficients do not necessarily lead to better smaller error detection probability.

2.3 Applications

For different applications and data types, the watermarking design requirements vary. In this section, several popular watermarking applications and their design requirements are introduced.

2.3.1 Copyright protection

The most common application of watermarking is the copyright protection. In this case, a watermark is used to identify the content ownership. For documents and images, this can be done through superposing visible watermarks in the form of patterns or texts to the host data. However, digital invisible watermarking is a better choice for its imperceptibility. In this case, the watermark is required to be robust. The distortions it has to resist are data type dependent. Detection reliability is also a main issue.

2.3.2 Content authentication

Content authentication is related to the integrity control of the data content. That is, the modification of the data content is indicated by the distorted watermark. For such applications, the watermark should tolerate minor modifications which do not alter the meaning of the data

content such as noise addition and compression errors. Therefore, such applications prefer the semi-fragile watermarks to the fragile ones. In essence, how much distortions the watermark should tolerate depends on the application purpose. For example, one requirement of such applications is that the locality of the distorted signals should be identified. Therefore, the watermark has to be spread all over the media content and thus the watermark payload requirement is larger than that for the copyright protection purpose.

2.3.3 Carrying Side Information

Watermarks can be used to carry side information as transmission over channels. In this subsection, we illustrate several practical examples.

2.3.3.1 Error Resilience

In a wireless multimedia transmission environment, strong multipath fading effects and other interferences degrade drastically the quality of transmitted data. Particularly, it is well known that the compressed multimedia bit streams are sensitive to data errors. This is also true for certain wired network environments. As a result, there is an urgent need for higher layer applications to offer good error resilience ability.

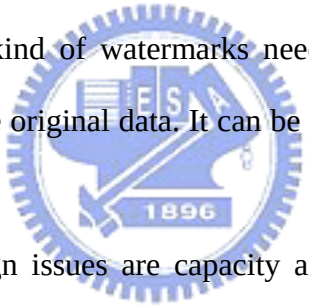
Many error resilience techniques for video/audio transmission have been suggested. Typically, an error concealment mechanism is effective if transmission errors can be correctly identified in the received data. It would be very helpful if additional information can be used to improve error detection ability while little penalty is paid in bit rate and/or compression efficiency.

Watermarking technologies provide the ability to hide data in the given digital signal without interfering human perception. Furthermore, a well-designed watermarking embedding scheme should not increase the total bit rate significantly. No extra header bits are needed either. Therefore, it seems to be a useful idea to use watermarks for error control purposes.

Watermarks can be parity check codes or any other type of error control/protection coding scheme selected according to the application requirements. For example, to combat burst errors in certain channels, watermark can be RS codes.

For image applications, parity check bits can be used to protect picture headers and image content. For video applications, parity check bits can be generated from macroblocks, motion vectors, GOP headers, GOB headers, etc. It is preferred that the information bits derived from one picture frame are embedded into the other frames. In case a frame is corrupted, its watermark (parity check codes) can be retrieved from the other frames. In [5], the watermarking scheme is used to refine images. In [6][7], such idea is realized to protect H.263 bitstreams.

Since the watermark is not used to prove ownership or to protect copyright, we expect few intentional attacks. This kind of watermarks need not be irremovable. Its function is helping detecting/protecting the original data. It can be discarded when no further decoding is expected.



The most important design issues are capacity and robustness against channel errors. How to increase the capacity of embedded information may be of highest priority. The amount and strength of the watermark can be designed with the help of a perceptual model. Another requirement is that the original digital data should not be required for extracting the watermark.

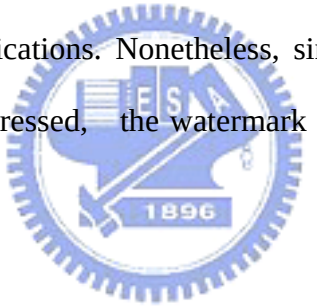
2.3.3.2 Data Hiding for Video-in-Video and Speech-in-Video

Watermarking can be used to convey supplementary information. It is proposed in [8][9]. The supplementary information can be video, audio, or speech. For example, speech and video are used as watermarks. They are compressed and embedded in the host video signal. The watermark can be public or secret. If it is secret, it can be used to control the access level for different viewers. In the speech-in-video case, the application can be several language streams

embedded in the video for different viewers. In the video-in-video case, the supplementary video can be a small window displayed next to the host video without sending a separate data stream.

The advantage of using watermarks in this type of application is that neither extra header nor separate data stream is needed. Thus, we can avoid the loss of information when the file format is converted. No separate stream eliminates the needs for extra transmission channels. Also, the synchronization problem between the host signal and the supplementary signal is solved when we edit the host signal streams.

However, such applications are useful only when the data hiding capacity is very high. Otherwise, the quality of the supplementary signal is low and may be useless. Thus, the most important design issue of such watermarking scheme is capacity. Few intentional attacks are expected for this type of applications. Nonetheless, since the host digital signal after data embedding may often be compressed, the watermark should survive common compression methods.



2.3.3.3 Annotation

The watermark can be also used as annotations for content management. For example, it can be taken as the metadata for digital rights management (DRM) as the specifications defined in MPEG-21. MPEG-21 standard is concerned with defining digital item declaration, content representation and usage, and terminal/network interfaces. By incorporating the MPEG-7 description scheme and the watermarking technique for managing the content identification number, a video archive system may be realized as shown in [10].

2.3.4 Network Quality Monitoring

For the multimedia applications over band-limited transmission environment, watermarking

enables a chance for the quality assessment of channel condition estimation and codec optimization. Since the watermark is inseparable from the media content, the channel condition can be estimated by comparing the extracted watermark with the original one on the decoder side, rather than referencing to the original multimedia content. As a result, the content provider can judge the quality of service, and thus the users are billed based on the perceived quality of viewing experiences. One example of such system for video applications is developed in [11].

2.3.5 Executable watermarks

In [12], the idea of executable watermarks is described. The authors suggest that it is possible to embed into a digital signal its associated applet running at the receiver. This type of watermarks is quite interesting and may be useful for certain applications. Since the watermark is always contained in the host digital signal, the receiver does not have to look around for an appropriate applet to run the host digital signal. No extra header bits are needed either. However, this may be viewed a special case of meta-data where the meta-data are now executable codes.